ENGELBERG CENTER for Health Care Reform at BROOKINGS

Brookings Roundtable on Active Medical Product Surveillance:

Developing Methods for Timely and Frequent Data Accrual in Vaccine Safety Surveillance

---

Katherine Yih, Epidemiologist and Lecturer, Harvard Pilgrim Health Care Institute and Harvard Medical School

May 29, 2013

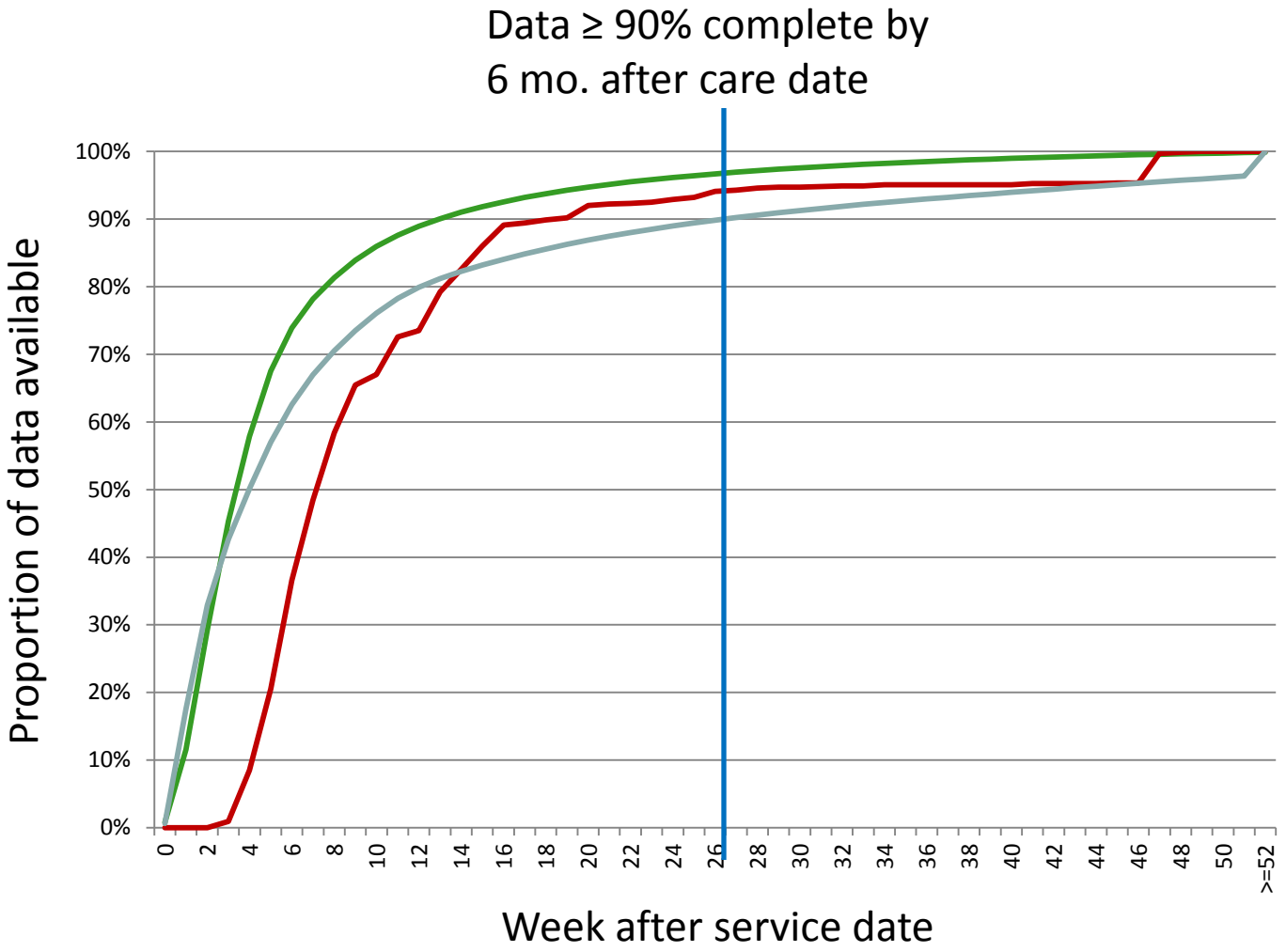
Using Freshest Feasible Data for Medical Product Safety Surveillance in Mini-Sentinel: Potential and Challenges

W. Katherine Yih, PhD, MPH

Harvard Pilgrim Health Care Institute and
Harvard Medical School

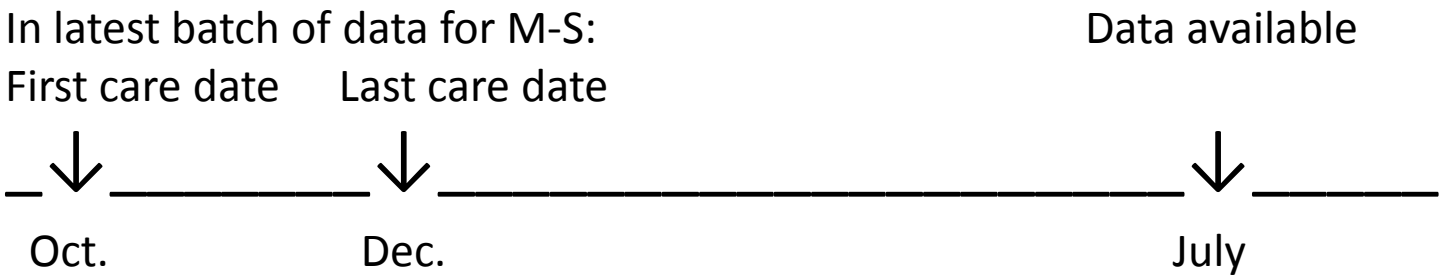
January 31, 2013

Inpatient claims data lag, 3 data partners



Mini-Sentinel data are relatively complete

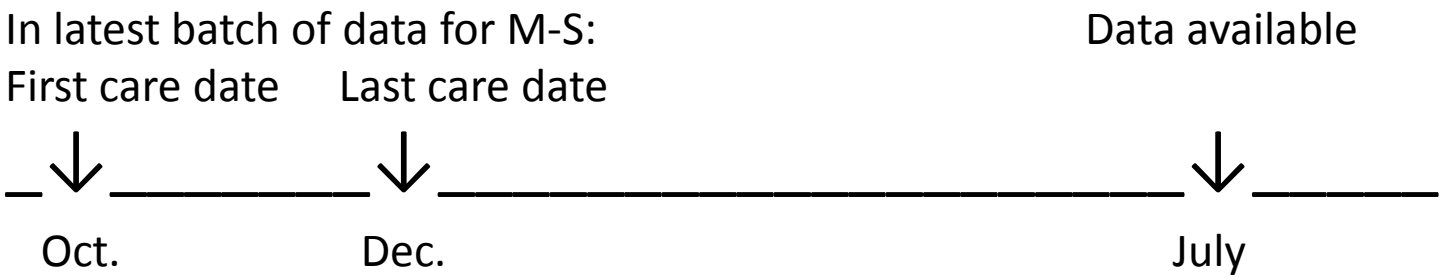
- Data updated on quarterly basis
- Typical example of timing:



- The most recent data typically 6-9 months old

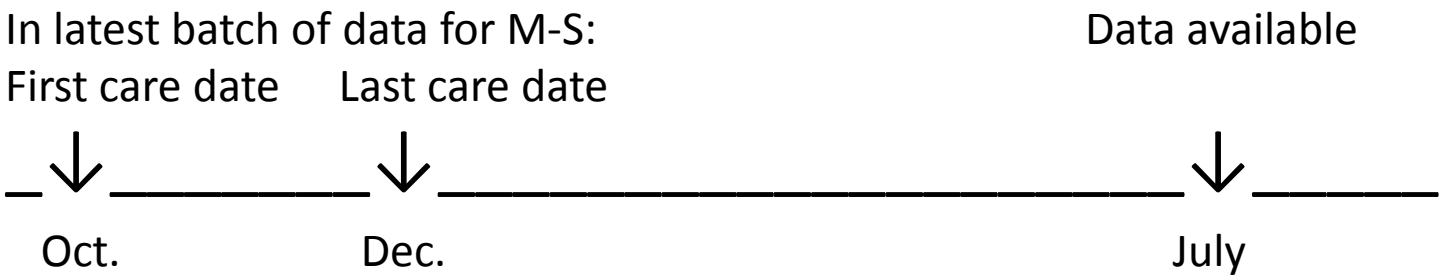
Advantage of mature (less fresh) data

- PRO: data more complete and settled



Pros and cons of mature (less fresh) data

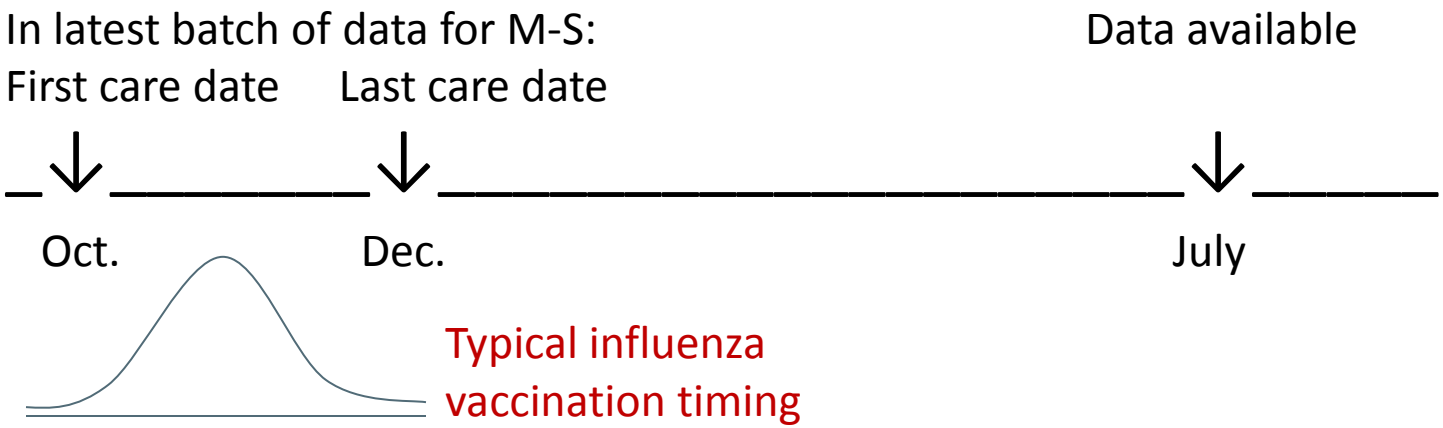
- PRO: data more complete and settled



- CON: signal detection delayed

Pros and cons of mature (less fresh) data

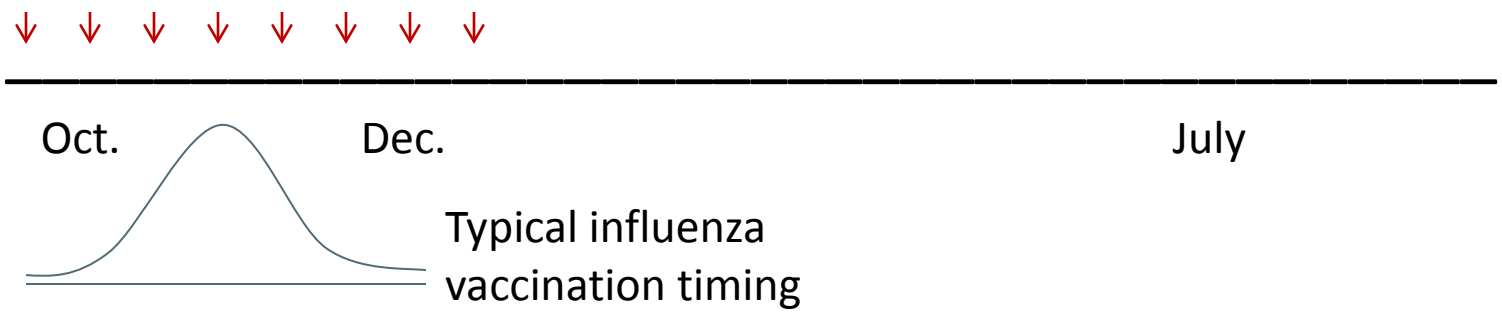
- PRO: data more complete and settled



- CON: signal detection delayed
  Especially problematic for influenza vaccine safety monitoring

Challenges of influenza vaccine safety monitoring

Influenza vaccination period relatively short, so data must be available soon after exposure to find safety problems in time to make a difference



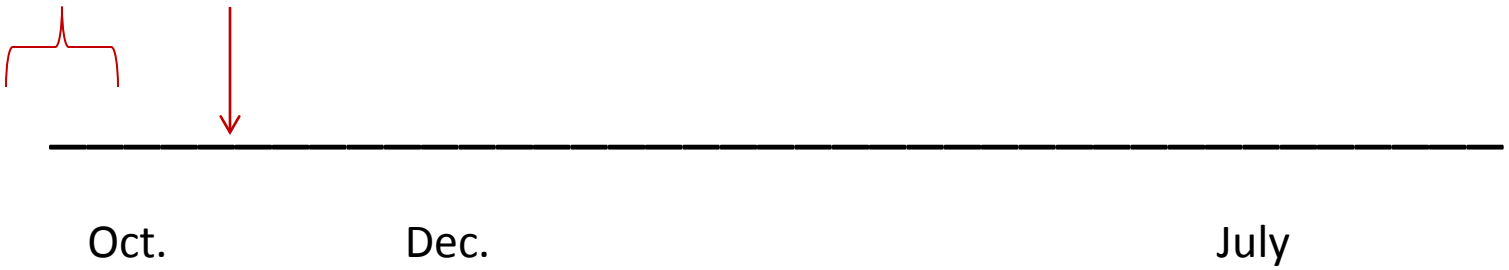
Challenges of influenza vaccine safety monitoring

1. Getting fresher and frequent data
2. Adjusting for incomplete data
3. Dealing with flux in the data over time

1. Getting fresher and frequent data

Freshest feasible data source is refreshed monthly

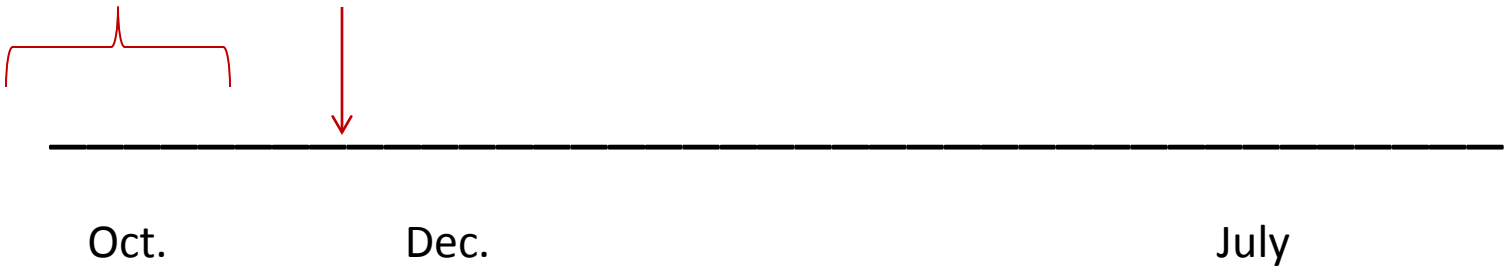
- Available toward end of following calendar month (data through Sept. available late Oct., etc.)
- More timely than M-S Distributed Dataset



1. Getting fresher and frequent data

Freshest feasible data source is refreshed monthly

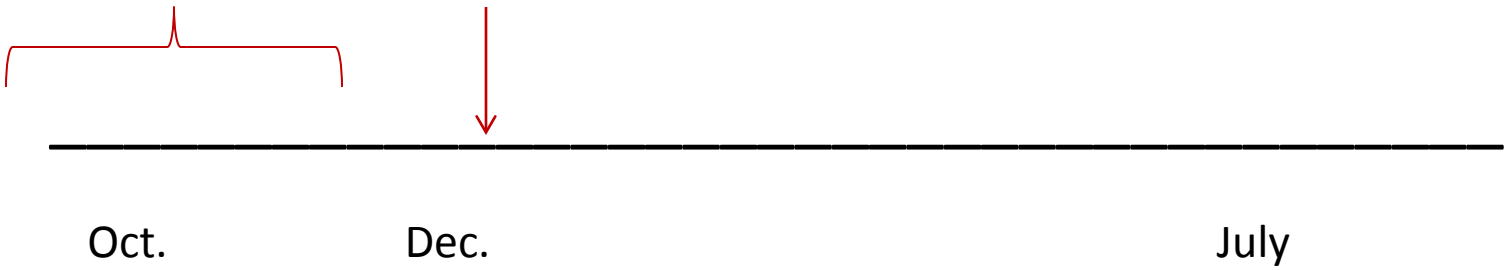
- Available toward end of following calendar month (data through Sept. available late Oct., etc.)
- More timely than M-S Distributed Dataset



1. Getting fresher and frequent data

Freshest feasible data source is refreshed monthly

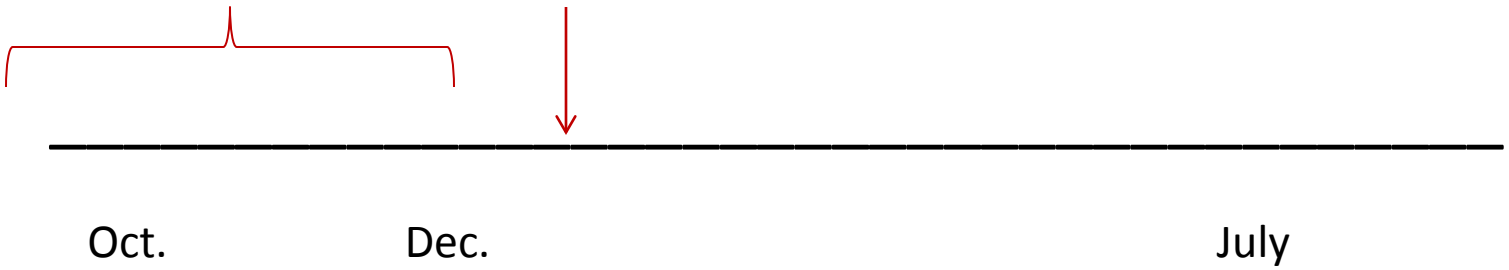
- Available toward end of following calendar month (data through Sept. available late Oct., etc.)
- More timely than M-S Distributed Dataset



1. Getting fresher and frequent data

Freshest feasible data source is refreshed monthly

- Available toward end of following calendar month (data through Sept. available late Oct., etc.)
- More timely than M-S Distributed Dataset



Files to be created for influenza vaccine safety monitoring

SDFs

Sequential Data Files (SDFs)

- Patient-level data, kept by data partners
- Population = persons with medical claim on or after 9/1/2012

SCFs

Sequential Case Files (SCFs)

- Patient-level data, kept by data partners
- Population = persons per current SDFs with health outcome of interest following influenza vaccination

SAFs

Sequential Analysis Files (SAFs)

- Aggregate data, sent to M-S Operations Center for analysis
- Vaccination population: vaccination per current SDFs
- Cases population: cases per all SCF versions

Expected timing of data refreshes and analyses

- Monthly but unsynchronized data refreshes by data partners
- Biweekly analyses by Operations Center (in weeks in red)

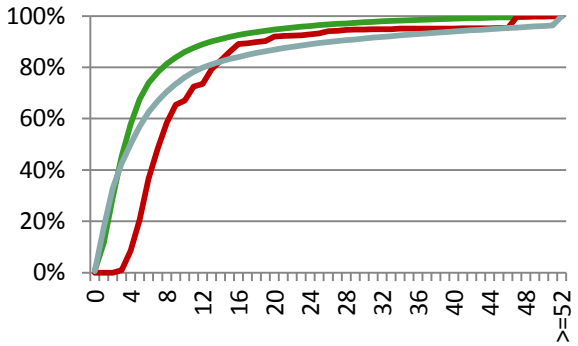
Week	1	**2**	3	**4**	5	**6**	7	**8**	9
DP1	SDF	**SAF**			SDF	**SAF**			SDF
DP2		SDF	**SAF…**	→		SDF	**SAF…**	→	
DP3			SDF	**SAF**			SDF	**SAF**	
Analysis		yes		yes		yes		yes	

2. Adjusting for incomplete data

Two kinds of "incompleteness"

A. Post-vaccination follow-up interval not fully elapsed

B. Lag in data availability →



To avoid bias, both must be taken into account.

PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2011; **20**: 583–590
Published online 29 April 2011 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.2133

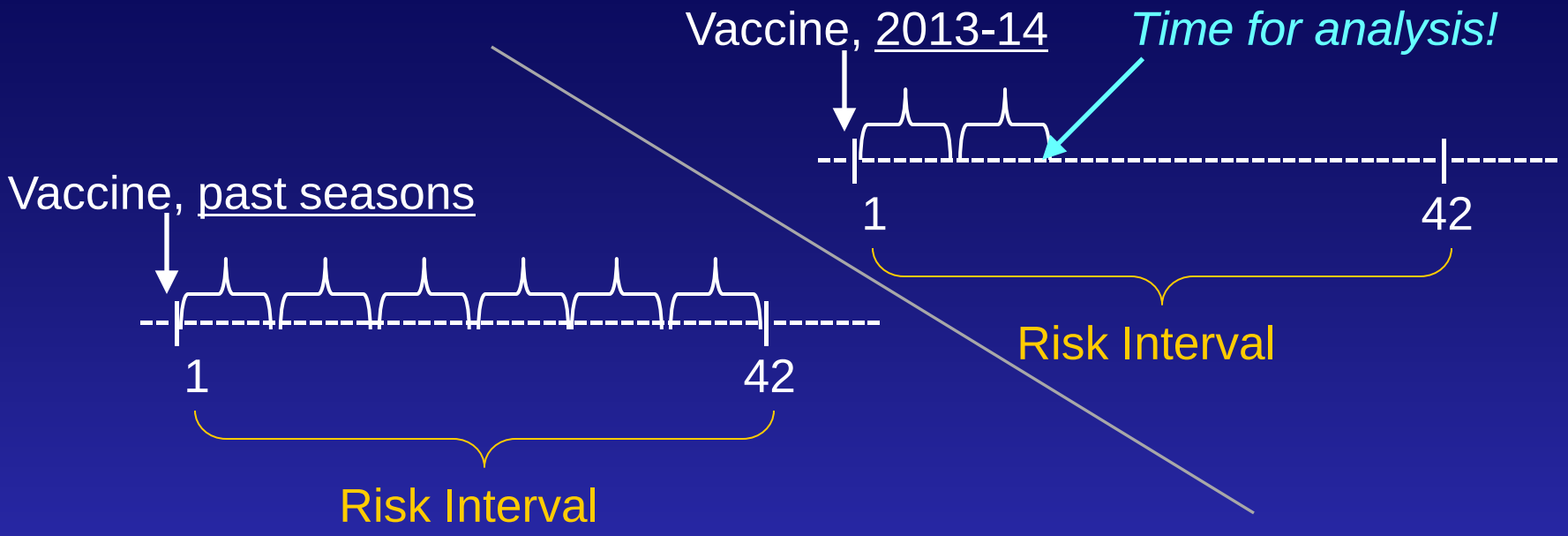
ORIGINAL REPORT

Near real-time vaccine safety surveillance with partially accrued data†

Sharon K. Greene[1]*, Martin Kulldorff[1], Ruihua Yin[1], W. Katherine Yih[1], Tracy A. Lieu[1], Eric S. Weintraub[2] and Grace M. Lee[1,3]

[1]*Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA*
[2]*Immunization Safety Office, Centers for Disease Control and Prevention, Atlanta, GA, USA*
[3]*Division of Infectious Diseases and Department of Laboratory Medicine, Children's Hospital Boston, Boston, MA, USA*

Issue A: Ongoing Risk Interval



- **Multiply expected events by proportion of risk interval elapsed, e.g.:**

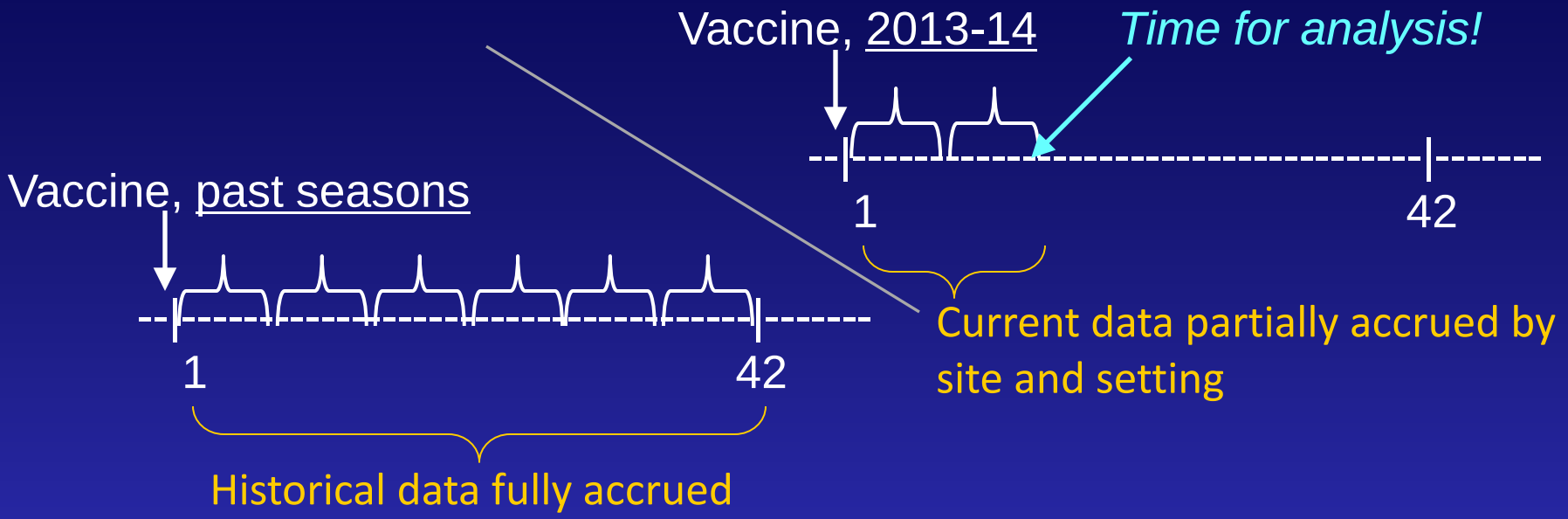
  10 expected events in 6 weeks following vaccination

  *** 1/3 interval elapsed**

  = 3.3 expected events in 2 weeks following vaccination

- **If don't adjust, expected events too high, biasing away from detecting safety problem**

Issue B: Late-Arriving Adverse Events



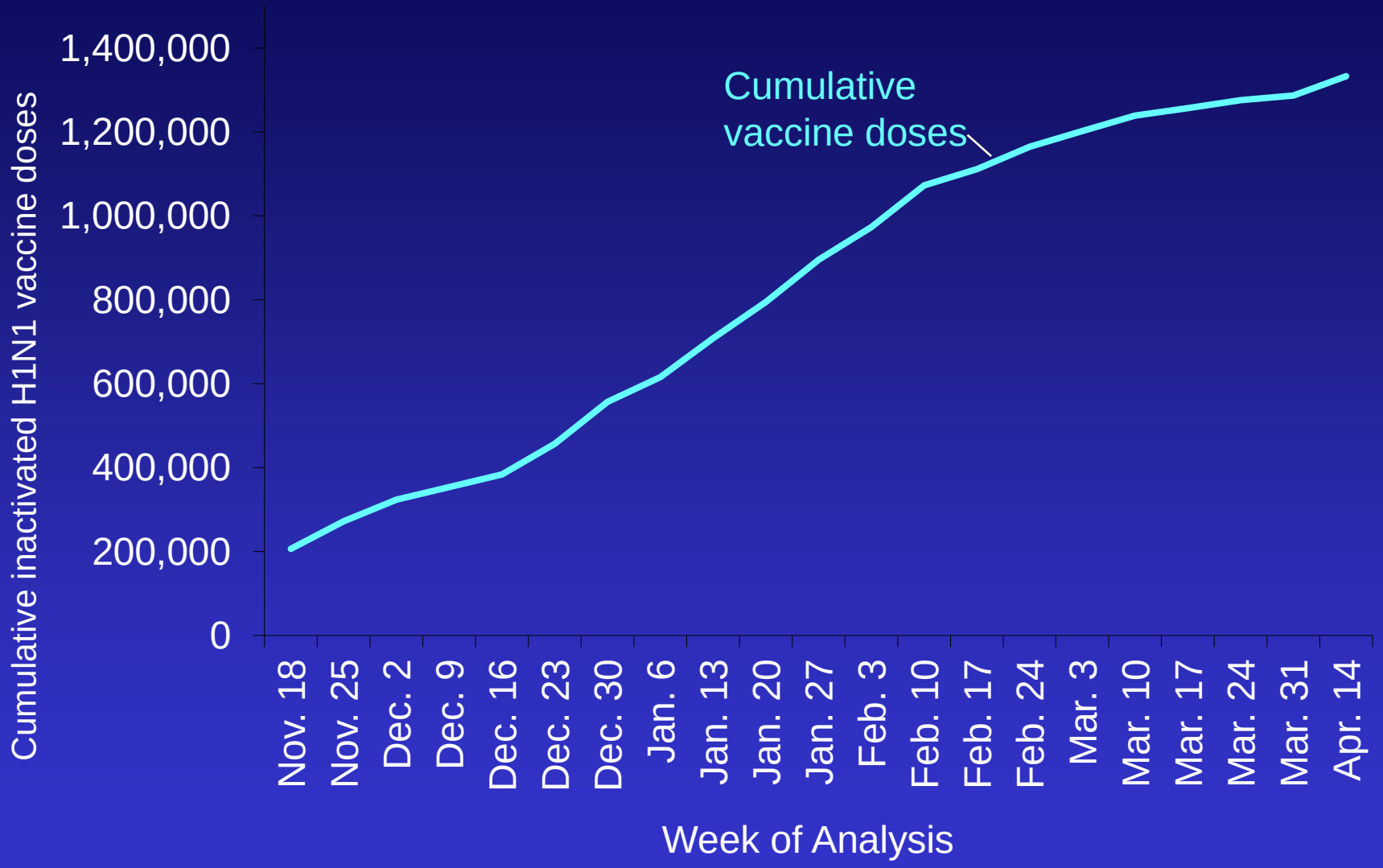
- **Further multiply expected number of events, e.g.:**

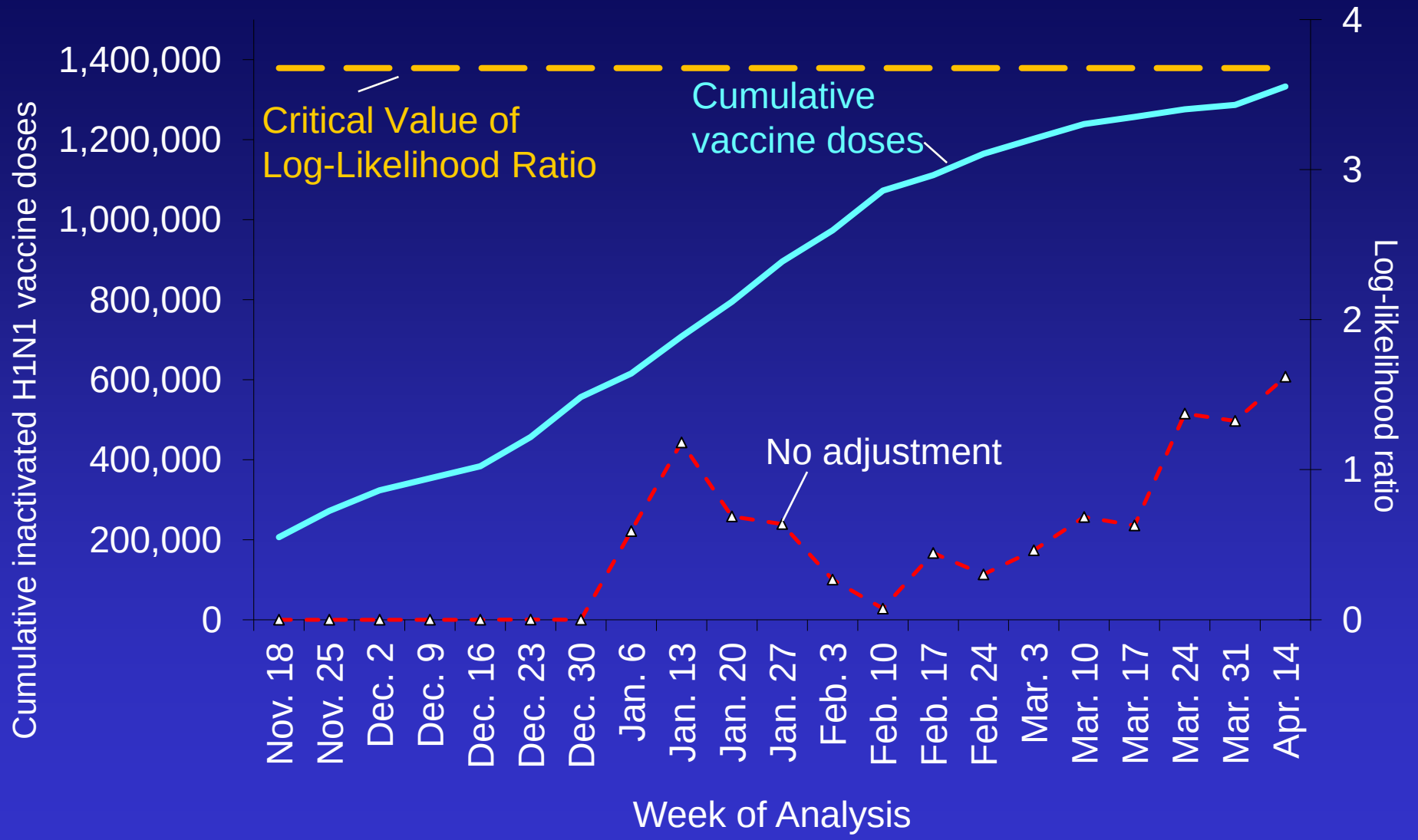
3.3 expected events in 2 weeks following vaccination

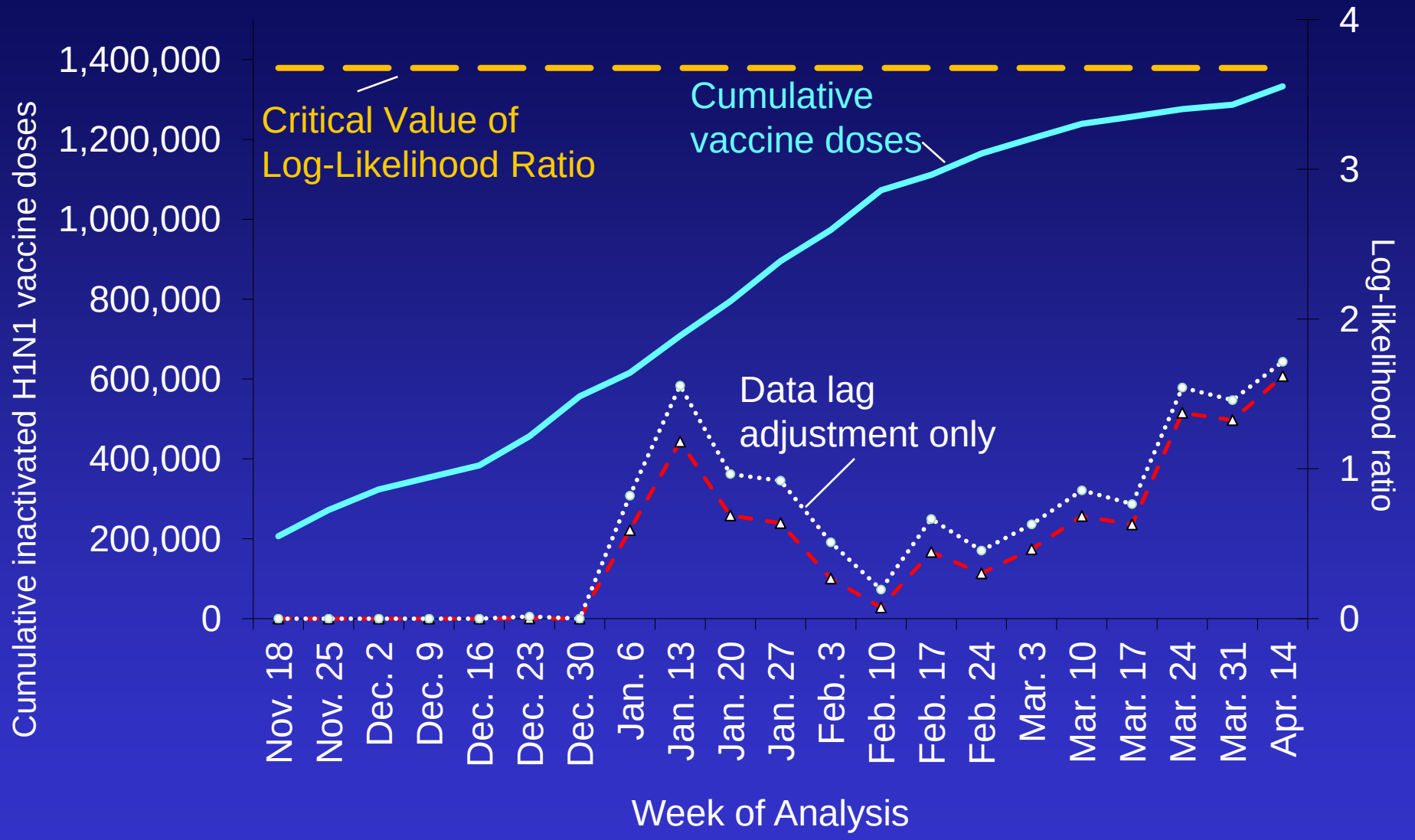
*** ((25% events expected in inpatient * 5% inpatient data accrued)**

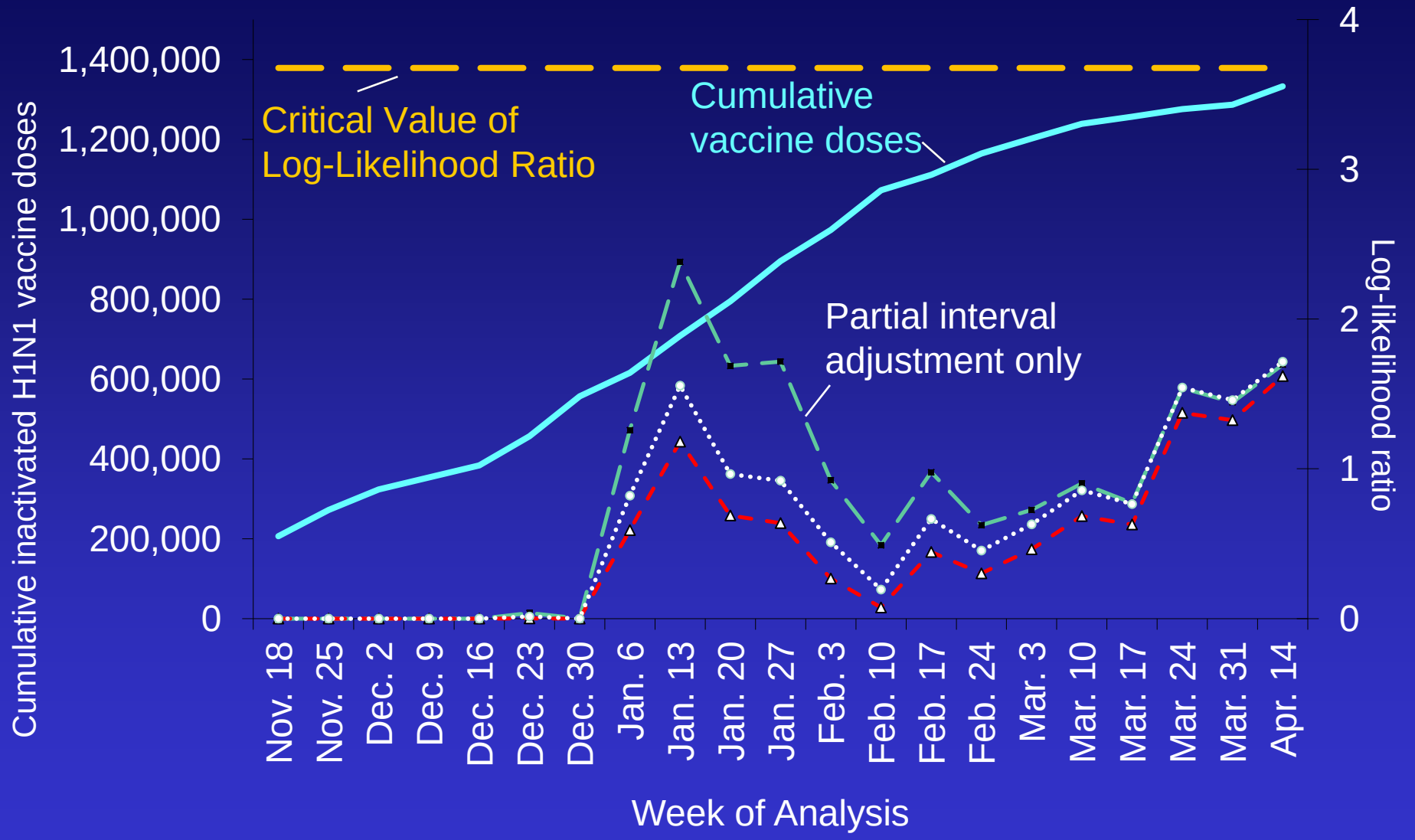
+ (75% events expected in ED * 60% ED data accrued))

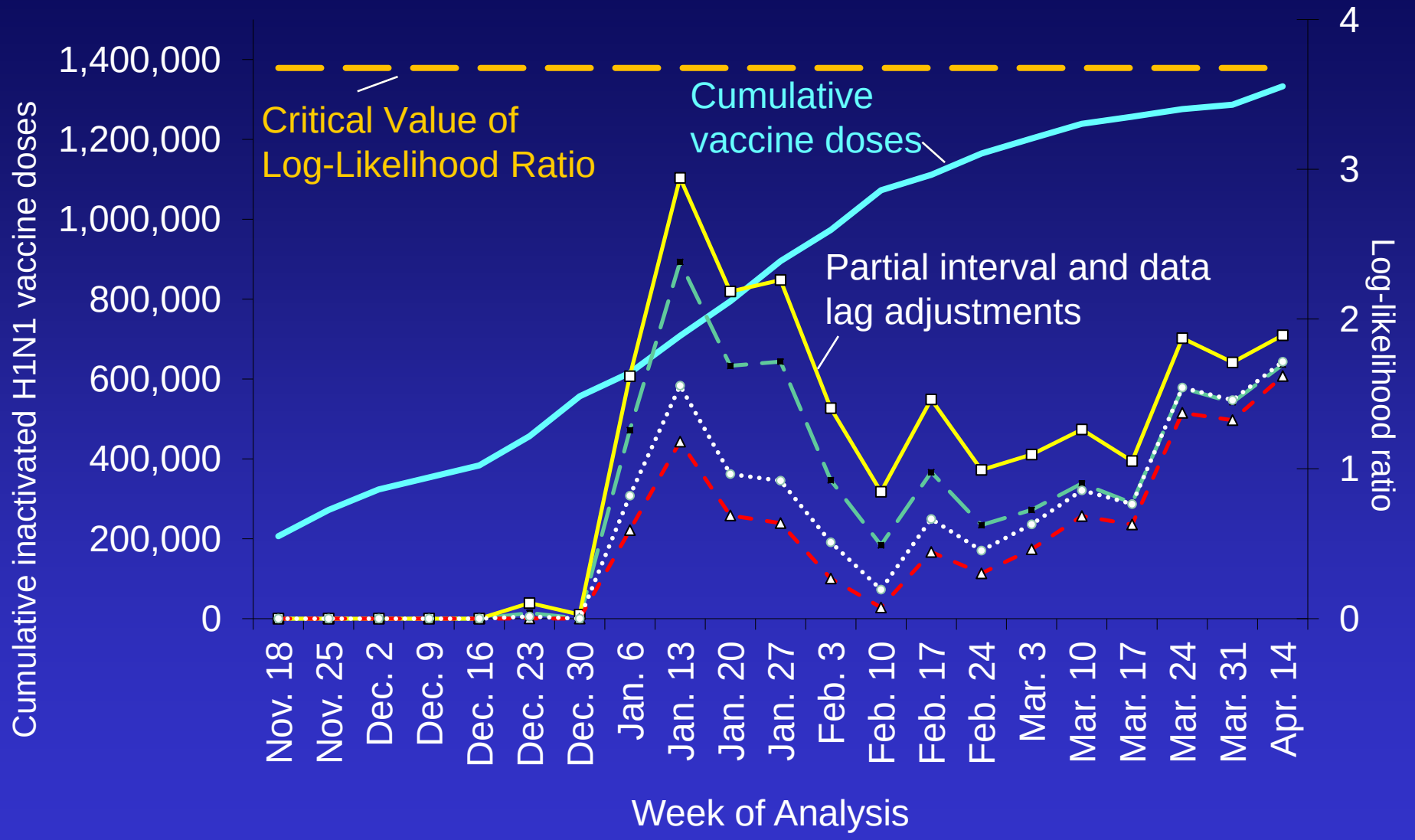
= 1.5 expected events in 2 weeks following vaccination, adjusted for data lags


Cumulative inactivated H1N1 vaccine doses
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
Cumulative vaccine doses
Nov. 18
Nov. 25
Dec. 2
Dec. 9
Dec. 16
Dec. 23
Dec. 30
Jan. 6
Jan. 13
Jan. 20
Jan. 27
Feb. 3
Feb. 10
Feb. 17
Feb. 24
Mar. 3
Mar. 10
Mar. 17
Mar. 24
Mar. 31
Apr. 14
Week of Analysis


Cumulative inactivated H1N1 vaccine doses
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
Log-likelihood ratio
4
3
2
1
0
Critical Value of
Log-Likelihood Ratio
Cumulative
vaccine doses
No adjustment
Nov. 18
Nov. 25
Dec. 2
Dec. 9
Dec. 16
Dec. 23
Dec. 30
Jan. 6
Jan. 13
Jan. 20
Jan. 27
Feb. 3
Feb. 10
Feb. 17
Feb. 24
Mar. 3
Mar. 10
Mar. 17
Mar. 24
Mar. 31
Apr. 14
Week of Analysis


Cumulative inactivated H1N1 vaccine doses
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
Log-likelihood ratio
4
3
2
1
0
Critical Value of
Log-Likelihood Ratio
Cumulative
vaccine doses
Data lag
adjustment only
Nov. 18
Nov. 25
Dec. 2
Dec. 9
Dec. 16
Dec. 23
Dec. 30
Jan. 6
Jan. 13
Jan. 20
Jan. 27
Feb. 3
Feb. 10
Feb. 17
Feb. 24
Mar. 3
Mar. 10
Mar. 17
Mar. 24
Mar. 31
Apr. 14
Week of Analysis


Cumulative inactivated H1N1 vaccine doses
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
Critical Value of
Log-Likelihood Ratio
Cumulative
vaccine doses
Partial interval
adjustment only
4
3
2
1
0
Log-likelihood ratio
Nov. 18
Nov. 25
Dec. 2
Dec. 9
Dec. 16
Dec. 23
Dec. 30
Jan. 6
Jan. 13
Jan. 20
Jan. 27
Feb. 3
Feb. 10
Feb. 17
Feb. 24
Mar. 3
Mar. 10
Mar. 17
Mar. 24
Mar. 31
Apr. 14
Week of Analysis


Cumulative inactivated H1N1 vaccine doses
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
Critical Value of
Log-Likelihood Ratio
Cumulative
vaccine doses
Partial interval and data
lag adjustments
4
3
2
1
0
Log-likelihood ratio
Nov. 18
Nov. 25
Dec. 2
Dec. 9
Dec. 16
Dec. 23
Dec. 30
Jan. 6
Jan. 13
Jan. 20
Jan. 27
Feb. 3
Feb. 10
Feb. 17
Feb. 24
Mar. 3
Mar. 10
Mar. 17
Mar. 24
Mar. 31
Apr. 14
Week of Analysis

3. Dealing with flux in the data over time

General kinds of flux:

1. Gain of cases (expected!)
2. Loss of cases
3. Reappearance of cases that had been lost
4. Changes in characteristics important to analysis, e.g. age group, dx date, medical setting

To maintain integrity of statistical testing:

Freeze data and results from prior sequential analyses

Flux in seizure cases* between two most recent data refreshes, 2012-13 pilot

	DP1	DP2	DP3
Time span between the two refreshes	5 mo.	1 mo.	2 mo.
In most recent data refresh (cumulative, no cases removed even if they disappeared since last refresh)	60	14	207
New since previous refresh	57 (95%)	6 (43%)	86 (42%)
Retained from previous refresh, no changes	0	8 (57%)	121 (58%)
Retained from previous refresh, change in characteristics	0	0	0
Lost since previous refresh	3 (5%)	0	0

* in 6-23 mo. olds in the 42 d after influenza vaccination

Conclusion

- PROS of using fresher data
  - Gain in timeliness ~5-8 mo.
  - Necessary for influenza vaccine safety monitoring
- CONS of using fresher data
  - Some loss of accuracy despite adjustments for data incompleteness and flux
  - Takes extra effort to produce these data—more frequent refreshes, different source files, special file structures
  - Each product needs a separate extract
- We *can* use fresher data, but probably not worthwhile to do so on routine basis